

## DICTATION SYSTEM

**Publication number:** WO9729578

**Publication date:** 1997-08-14

**Inventor:** BOELDL HERBERT

**Applicant:** PHILIPS ELECTRONICS NV (NL); PHILIPS NORDEN AB (SE)

**Classification:**

- international: **G06F3/16; G10L19/00; G06F3/16; G10L19/00;** (IPC1-7): H04M

- European: G10L19/00U

**Application number:** WO1997IB00048 19970124

**Priority number(s):** EP19960200328 19960212

**Also published as:**



WO9729578 (A3)  
US6182043 (B1)  
CN1185854 (A)  
DE69721404T (T2)  
CN1119797C (C)

**Cited documents:**

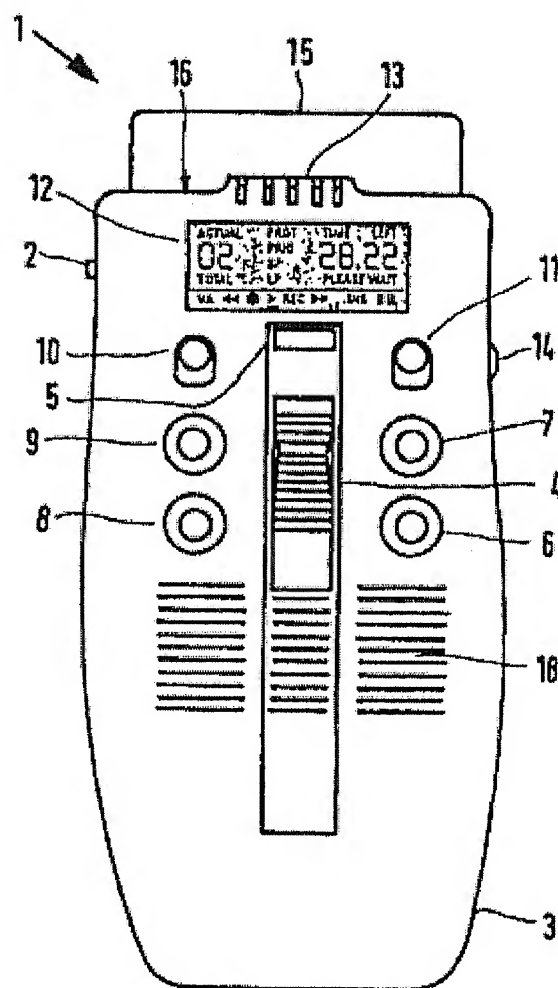


US5481645  
WO9528702  
JP8083099

**Report a data error here**

### Abstract of WO9729578

A dictation system is disclosed comprising a hand-held dictation device (1) for storing a speech signal in memory means (15, 20), the device comprising data compression means (30) for data compressing the speech signal into a data compressed speech signal and storing means for storing the data compressed speech signal in the memory means. The data compression means (30) are adapted to carry out a data compression step on the speech signal in one of at least two different data compression modes, the at least two different data compression modes resulting in different data compression ratios when applied to the same speech signal, the said at least two different data compression modes being selectable by a user. The data compression means (30) are further adapted to create data files (Bi) comprising portions of the data compressed speech signal, the data files comprising a header portion (HDR), the data compression means being also adapted to generate an identifier signal identifying the data compression mode selected and being adapted to store said identifier signal in said header portion.



Data supplied from the esp@cenet database - Worldwide



[12] 发明专利申请公开说明书

[21] 申请号 97190285.2

[43]公开日 1998 年 6 月 24 日

[11] 公开号 CN 1185854A

[22]申请日 97.1.24

[30]优先权

[32]96.2.12 [33]EP[31]96200328.1

[86]国际申请 PCT/IB97/00048 97.1.24

[87]国际公布 WO97/29578 英 97.8.14

[85]进入国家阶段日期 97.12.1

[71]申请人 飞利浦电子有限公司

地址 荷兰艾恩德霍芬

[72]发明人 H·博尔德尔

[74]专利代理机构 中国专利代理(香港)有限公司

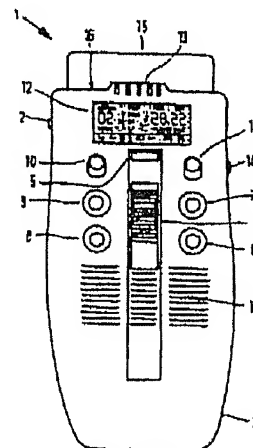
代理人 王 勇 王 岳

权利要求书 3 页 说明书 6 页 附图页数 3 页

[54]发明名称 口授系统

[57]摘要

此处公开的口授系统包括一种将语音信号存于存储部件(15, 20)的手持式口授装置(1), 该装置包含能将语音信号压缩成数据压缩语音信号的数据压缩部件(30)和存储数据压缩语音信号在存储器部件内的存储部件。数据压缩部件(30)被设计成以至少可用两种不同的数据压缩模式之一对语音信号进行数据压缩, 当用至少两种不同的数据压缩模式对同一语音信号进行压缩时将得到不同的数据压缩比。用户可选择至少所述两种不同的数据压缩模式。数据压缩部件(30)还适于生成包含数据压缩语音信号部分的数据文件(Bi), 数据文件包含一文件头(HDR)。数据压缩部件还适于生成用于标识被选数据压缩模式的标识信号并将该标识信号保存于文件头中。



# 权 利 要 求 书

1.一种口授系统, 包括一种用于将语音信号存于存储部件中的手持式口授装置, 该装置包含能将语音信号压缩成数据压缩语音信号的数据压缩部件, 和在存储部件中存储数据压缩语音信号的存储部件, 其特征  
5 在于数据压缩部件被设计成至少可用两种不同的数据压缩模式之一对语音信号进行数据压缩, 当对同一语音信号用至少两种不同的数据压缩模式进行压缩时会得到不同的数据压缩比, 用户可选择所述至少两种不同的数据压缩模式, 数据压缩部件还可生成包含数据压缩语音信号部分的数据文件, 该数据文件包含一个文件头, 数据压缩部件还适于生成用  
10 于标识所选择的数据压缩模式的标识信号并将其保存于文件头中。

2.权利要求1的口授系统, 其特征在于存储部件包括用于存储该数据文件的可拆卸式固态存储单元, 该固态存储单元带有连接部件, 用于将存储单元与手持式口授装置在机械和电气上相连。

3.权利要求2的口授系统, 其特征在于连接部件还被设计成在机械  
15 和电气上将该存储单元与PC机相连。

4.权利要求3的口授系统, 其特征在于连接部件被设计成在机械和电气上将该存储单元与该PC机的国际标准接口相连。

5.权利要求4的口授系统, 其特征在于所述接口就是PCMCIA接口。

20 6.权利要求2-5中任一项权利要求的口授系统, 其特征在于固态存储单元包括一种EEPROM。

7.权利要求2-5中的任一项权利要求的口授系统, 其特征在于固态存储单元包括快速可擦式存储单元。

8.权利要求1-5中任一项权利要求的口授系统, 其特征在于固态  
25 存储单元包括一个后备电池。

9.前述权利要求中任一项权利要求的口授系统, 其特征在于数据压缩部件被设计成至少可用两种不同的数据压缩模式之一对语音信号进行数据压缩处理, 所述至少两种数据压缩模式中至少有一种是带损耗的数据压缩模式。

30 10.用于前述任一项权利要求的口授系统中的手持式口授装置, 该装置包括将语音信号压缩成数据压缩语音信号的数据压缩部件, 其特征在

于, 该数据压缩部件被设计成至少能用两种不同的数据压缩模式之一对语音信号进行数据压缩处理, 该至少两种不同的数据压缩模式所得的压缩比是不同的, 用户可选择所说的至少两种不同的数据压缩模式, 该数据压缩部件还可生成包含数据压缩语音信号部分的数据文件, 该数据文件包括一个文件头, 该数据压缩部件还生成用于标识被选的数据压缩模式的标识信号并将该标识信号保存于所述文件头中。

11. 权利要求 10 的手持式口授装置, 其特征在于, 它具有与可拆卸式固态存储单元的连接部件在机械上和电气上相连的连接部件。

12. 权利要求 11 的手持式口授装置, 其特征在于连接部件是按照一种国际  
10 标准接口制造的。

13. 权利要求 12 的手持式口授装置, 其特征在于所述的接口是 PCMCIA 接口。

14. 权利要求 10 - 13 中任一项权利要求的手持式口授装置, 其特征在于数据压缩部件适于用至少两种不同的数据压缩模式之一对语音信号进行数据压缩, 所说的至少两种不同的数据压缩模式中至少有一种是带损耗的数据压缩模式。

15. 用于复读语音信息的复读装置, 其特征在于它包括用于对存储于存储部件中的数据压缩语音信号进行解压操作的数据解压部件, 所述数据压缩语音信号由至少两种不同数据压缩模式中的一种进行数据压缩而成, 对于相同语音信号, 用该至少两种不同的数据压缩模式进行数据压缩得到不同的数据压缩比, 该数据压缩语音信号作为数据文件存于存储部件中, 该数据文件包含数据压缩语音信号部分, 数据文件还包括存有标识信号的文件头, 该标识信号表示对语音信号进行数据压缩时所选的数据压缩模式, 数据解压部件能从文件头中检索到该标识信号, 并按此标识信号的要求用至少两种不同数据解压模式中的一种对文件中的数据压缩语音信息进行解压操作, 以便得到该语音信号的副本。

16. 权利要求 15 的复读装置, 其特征在于存储部件是可拆卸式固态存储单元, 该复读装置具有与该固态存储单元的连接部件在机械上、电气上相连的连接部件。

17. 权利要求 16 的复读装置, 其特征在于该连接部件是按照一种国际  
30 标准接口的要求制造的。

18. 权利要求 17 的复读装置, 其特征在于所述接口是 PCMCIA 接

口。

5 19.可拆卸式固态存储单元,包含数据压缩语音信号,所述数据压缩语音信号是采用至少两种不同的数据压缩模式中的一种进行数据压缩的,当对相同的语音信号用至少两种不同的数据压缩模式进行数据压缩时所得到的数据压缩比是不同的,该数据压缩语音信号以数据文件形式存于存储部件中,该数据文件包含数据压缩语音信号部分,该数据文件包括存有标识信号的文件头,该标识信号识别对语音信号进行数据压缩时所选的数据压缩模式。

10 20.权利要求 19 的固态存储单元,其特征在于它具有将该存储器与 PC 机在机械、电气上相连的连接部件。

21.权利要求 20 的固态存储单元,其特征在于该连接部件将该存储单元与 PC 机的一种国际标准接口在机械、电气上相连。

22.权利要求 21 的固态存储单元,其特征在于所述的接口是 PCMCIA 接口。

# 说明书

## 口授系统

本发明涉及一种口授系统 ( dictation system )，它包括一种将语音信号存储于存储部件中的手持式口授装置，此装置包含将语音信号压缩成压缩语音信号的数据压缩部件和在存储部件内存储压缩语音信号的存储部件。本发明还涉及一种在该口授系统中使用的口授装置、一种复读装置和一种可拆卸式固态存储单元。首段中定义的口授系统在本技术领域是众所周知的。

现有技术中的口授系统通过丢弃常见于语音信号中的无音周期来实现数据压缩。进而，可以保存一个指示信号，表示无音周期的长度和它在语音信号中的位置。在复读时，通过在所压缩的语音信号中所标记的地方插入相应的无语音周期，从而再生语音的副本。

本发明的目的在于提供一种改进的口授系统。根据本发明的口授系统，其特征在于数据压缩部件被设计成对语音信号至少以可用两种不同的数据压缩模式之一进行数据压缩，当用该至少两种不同的数据压缩模式对同一语音信号进行数据压缩时会得到不同的压缩比。用户可选择所说的至少两种不同的数据压缩模式。数据压缩部件还可进一步生成包含压缩语音信号部分的数据文件，该数据文件包含一个文件头。该数据压缩部件还生成用于标识所选的数据压缩模式的标识信号并将该标识信号存于所述文件头中。本发明基于下列认识：口授装置中的存储器容量有限。而存储部件中应最好存储数目不断增长的口授段，这在本领域中通过丢弃语音信号中的无语音周期来实现。利用更强有力的压缩技术可获得更高的压缩比。更具体地说，有损耗的压缩技术产生高数据压缩比。但是，较高的压缩比会导致数据展开时所恢复的语音信号的质量下降。本发明提出这样一种口授系统，它允许用户从两种或多种数据压缩模式中选择其一，手持式口授装置据此可对语音信号进行压缩。用户可在一个内存单元中可口授和存储的语音段数目与语音信号的重现质量之间作出权衡。假如用户希望内存部件中存储较多的语音段，他就应选择具有较高数据压缩比的数据压缩模式；假如用户更注重语音复读质量，他就应选择较低压缩比的数据压缩模式。

从属权利要求定义了口授系统，手持式口授装置，复读装置及可拆

卸式固态内存单元的较佳实施例。

参照下列实施例附图的说明，我们将弄清楚并阐明本发明的这些和其他方面。

图 1 表示手持式口授装置的一种实施例。

5 图 2 表示用手持式口授装置的存储卡的一种实施例。

图 3 表示手持式口授装置的电路图。

图 4 表示手持式口授装置中处理器生成的信号块序列。

图 5 是以桌面机或 PC 机形式实现的复读装置的一个实施例。

10 图 1 是手持式口授装置 1 的正面视图，开关 2 位于该装置外壳的侧面，外壳底部设有电池舱 3（没画出）。可以从外壳背面取放电池，壳的正面有滑动开关 4，用于在不同口授模式之间进行切换。该装置还有几个按钮：按钮 5 是录音按钮，按钮 6 是 LETTER（语句）按钮，按钮 7 是 MODE（模式）按钮，按钮 8 是 INSERT（插入）按钮，按钮 9 是 DELETE（删除）按钮。开关 10 是录音模式开关。开关 11 是灵敏度开关。装置 1 还提供显示有关口授的各种信息用的 LCD 显示器，例如，语音段的录制时间、剩余录制时间、录制模式、语音段数目等。

15 外壳内设有麦克风 13 和扬声器 18。壳的侧面装有音量控制旋钮 14。另外，该装置上侧有一插槽，用于插入存储卡 15。

20 存储卡 15 也示于图 2 中。存储卡 15 具有一个固态存储器 20 及多个与固态存储器 20 相连的电气端子 22。固态存储器 20 可以是，例如，EEPROM 或快速可擦式存储器。电气端子 22 在电气上能与 PC 机上的符合国际标准的 PCMCIA 接口相接。

25 图 3 表示装置 1 的电气结构及其与存储卡 15 的配合。装置 1 包括数字信号处理器 30，该处理器 30 具有与端子 34 相接的数字输入/输出端 32，而端子 34 在电气上又与插入槽 16 中的存储卡 15 的端子 22 相接。麦克风 13 连接处理器 30 的模拟输入端 36。若需要的话，可通过放大器 38 相接。处理器 30 还包括模拟输出端 40，它通过放大器 42 与扬声器 18 相接。图 3 中用参考数字 44 所指的各种旋钮和按钮与处理器 30 的控制输入端 46 相接。另外，处理器 30 的控制输出端 48 与用于控制显示器 12 上显示信息的显示控制单元 50 相接。

30 用户把存储卡 15 插入装置 1 的插槽 16 中，直到存储卡 15 的端子 22 与装置 1 的插槽中的端子 34 相接触。现在，存储卡就在电气和机械

上与装置 1 相接。

处理器 30 能通过输入端 36 接受模拟语音信号并通过 A/D 转换器将语音信号转换成数字语音信号。进而，通过用户的选择，处理器 30 能对该数字语音信号用至少两种不同的数据压缩步骤之一进行处理。假设  
5 处理器 30 能对该语音信号进行两种数据压缩处理，每种压缩处理步骤在对同一语音信号压缩后得到的压缩比是各不相同的。该数据压缩处理步骤可以是无损耗的。这意味着实际上没有丢失任何数据。且在解压数据时可得到与原有语音一模一样的语音信号。无损耗数据压缩方法的一个例子是线性预测编码，随后是对线性预测编码器的输出进行的  
10 Huffman 编码，数据压缩也可是带损耗的。一种带损耗数据压缩方法是本领域所熟知的分波段编码 (subband coding)，并应用于 DCC 数字式磁记录系统中。在带损耗数据压缩方法中，听不到的那部分语音信息实际上被丢弃了。数据解压就是恢复原始语音信号。既然在数据压缩时被丢弃的那部分语音是听不见的，那么在再现语音信号时用户听到的语音信号就和原有的一样。  
15

处理器 30 既能执行无损耗数据压缩处理步骤，也能执行带损耗的数据压缩处理步骤，这是处理器 30 能实现的两种不同数据压缩步骤。作为一种替代，处理器 30 能执行两种不同的无损耗数据压缩步骤，它们分别产生不同的数据压缩比。另外一种替代是，处理器 30 能对语音  
20 信号执行两种不同的无损耗数据压缩处理步骤，它们产生不同的数据压缩比。这儿给出后一种可能性的一个例子：与用于 DCC 中一样，给处理器 30 配备一个简单分波段编码器。可用简单的该分波段编码器是因为要求用于编码语音信号的分波段数目较少，因为语音信号的带宽比宽带音频信号小得多，只需较少数目的分波段，例如 5，而不是用于 DCC  
25 分波段编码器的 32。通过在对简化的分波段编码器进行位分配操作时改变其位池 (bitpool)，就可以简单化的分波段编码器得到不同压缩比。关于这方面的信息请参考列于本说明书末尾的文件 (1)，(2)，(3a)，(3b)。

当用户想往装置中灌录一段语音时，按一下 LETTER 按钮 6，此  
30 按钮表示用户想存储一段语音，另外用户可按动 MODE 按钮 7，以便选择不同的模式。如该语音段是否具有 (高) 优先权，或该语音段是否受保护以免被覆盖。随后用户启动按钮 10 来选择录制模式。选择录制



模式意味着用户选择了数据压缩模式。若用户想让录制音质较好，他/她就选择产生最低数据压缩比的数据压缩模式，这样存储器 20 将存储这一语音段的更多信息用于所述口授，结果是存储器能存储的语音段数目更少。如果用户想在存储器存储尽可能多的语音段，他/她就选择具有较高数据压缩比的数据压缩模式，结果是存储的语音段的音质较差。

所压缩的信息包括在信息块（或“文件”） $B_i$ 、 $B_{i+1}$ 、 $B_{i+2}$ ，... 中。如图 4 所示，每一信息块  $B_i$  具有一信息头（用 HDR 表示），和信息体（用 IP 表示）。另外，在信息头中存有一标识信号，信号块的头部 HDR 中的标识信号标识对语音信号的压缩模式，以便生成存储在同一个信号块的信息体 IP 部分的数据压缩信息。信号块的序列号提供给处理器 30 的数字输出端 32，然后存于存储卡 15 的存储器 20 中。

这里应当指出，处理器 30 能生成尽可能长的信号块用来完整记录每一语音段的信息。处理器 30 也能产生固定长度的信号块，这样，一个语音段的数据压缩信息就存储在由处理器 30 产生的多个连续的信号块中。

如果用户想得到存储在存储器 20 中的语音段，处理器 30 能恢复存储器 20 中的压缩数据信息并对存于存储器 20 中的压缩数据信息执行数据解压操作。显然，数据解压操作就是在口授阶段执行的数据压缩操作的逆操作。在后面用图 5 说明复读装置的实施例时，将进一步解释处理器 30 执行的数据解压操作。得到语音信号副本后，经处理器进行数/模转换，供给输出端 40，用于扬声器 18 重放。

为将存储于存储卡 15 上存储器 20 中的语音段复读出来，将装置 1 上的存储卡 15 拔出，插入桌面的复读装置 50 中，见图 5。复读装置 52 包含一数字信号处理器 53，它具有数字输入端 54，与端子 56 相连，当将存储卡插入该装置 52 中的插槽（没画出）时，该端子 56 在电气上和存储卡 15 的管脚 22 相连。扬声器 58 经放大器 62 与处理器 53 的模拟输出端 60 相连。处理器 53 还包括与显示器 68 上控制显示信息的显示控制单元 66 相连的控制输出端 64。键盘 70 与处理器 53 的控制输入端 72 相连。

用户将存储卡 15 插入复读装置 52 的插槽（没画出）中。直到存储卡 15 的端子 22 与复读装置 52 的插槽中的端子 56 相接触，现在存储卡在机械和电气上就与装置 52 相连。

按动键盘 70 上的 RETRIEVE (恢复) 键, 存储卡 15 上存储器 20 中的信息就被读出, 并保存在数字信号处理器 53 的内部存储器中。处理器对从存储卡中恢复的信息能执行至少两种不同的数据解压操作。显然, 处理器 53 执行的解压模式正好与处理器 30 在口授阶段执行的压缩模式相反。处理器 53 从信号块头部 HDR 中检索到标识信号, 并按此标识信号执行数据解压操作, 结果得到数字化语音信号的副本。

处理器 53 进一步将数字语音信号副本进行数/模转换转换成模拟语音信号, 并将模拟语音信号经输出端 60 供给扬声器 58, 这样打字员或别人就能听到需要复读的语音信号。

打字员用键盘 70 打出由扬声器播出的语音段, 于是就得到了该语音段的打字形式。

复读装置 52 的另一实施例, 采用具有充分大的存储容量的个人计算机时, 该装置可配上语音识别算法, 使该装置在对语音信号识别后从语音信号生成字符文件, 借助显示器 68 能见到字符文件的内容。这样打字员就能一边阅读显示屏幕 68 上的文本检查错误, 一边监听从扬声器 58 出来的语音, 同时还可通过键盘 70 改正文本中的错误。

前面介绍了一个无损耗数据压缩方法的例子, 即线性预测编码, 后随 Huffman 编码。不言而喻, 处理器 53 肯定能执行相应的 Huffman 解码, 后随相应的线性预测解码, 用于再生原语音信号。

前面也介绍了一种带损耗的数据压缩方法的实例, 即分波段编码。不言而喻, 处理器 53 应能执行相应的分波段解码操作, 用于再生成原语音信号的副本。

尽管本说明描述了本发明及其优选实施方案, 应当相信, 并不局限于上述这些例子。因此, 本领域的技术人员可在不超出本发明权利要求规定的范围内作各种明显的改动。进一步讲, 本发明在于这儿公开的特性所蕴含的每个新特性或各种特性的组合。

相关文档:

( 1 ) 欧洲专利申请号: 402,973(PHN13.241)

( 2 ) 欧洲专利申请号: 400,755(PHQ89.018A)

( 3a ) 欧洲专利申请号: 457,390(PHN13.328)

( 3b ) 欧洲专利申请号: 457,391(PHN13.329)

# 说明书附图

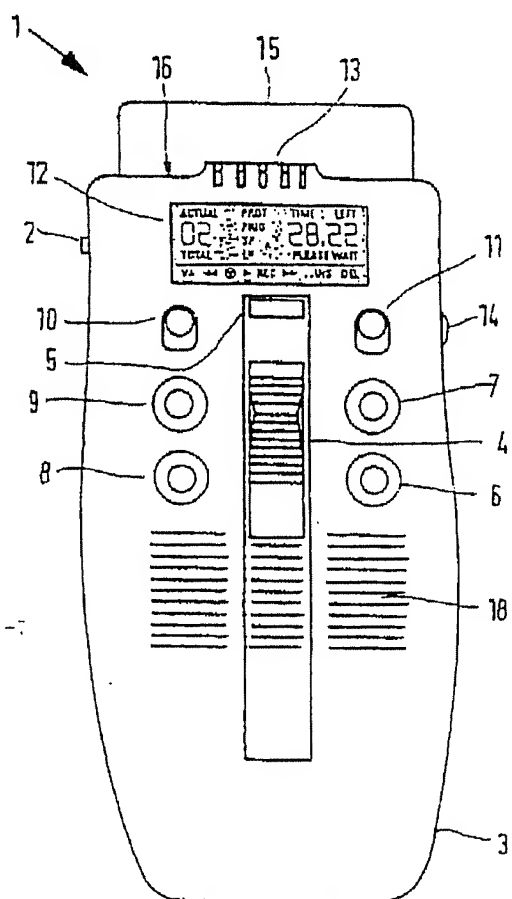


图 1

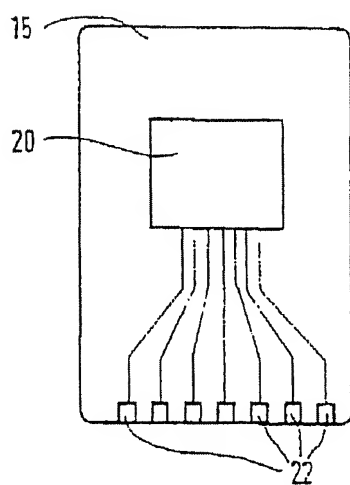


图 2

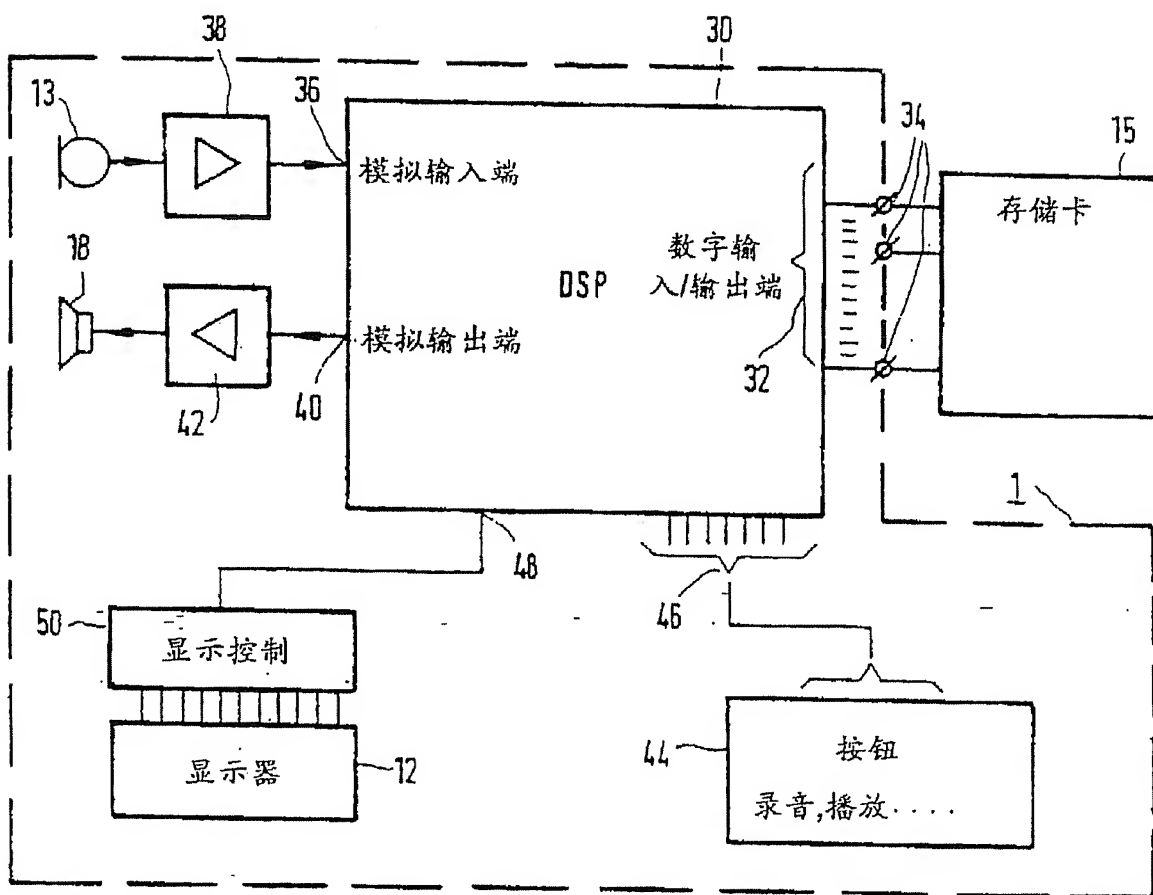


图 3

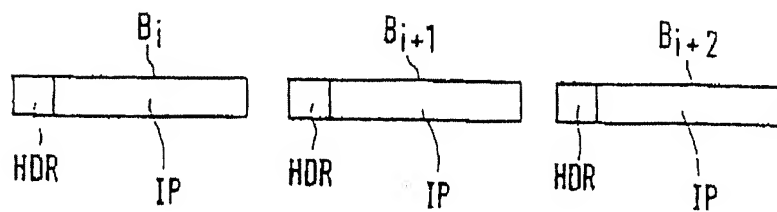


图 4

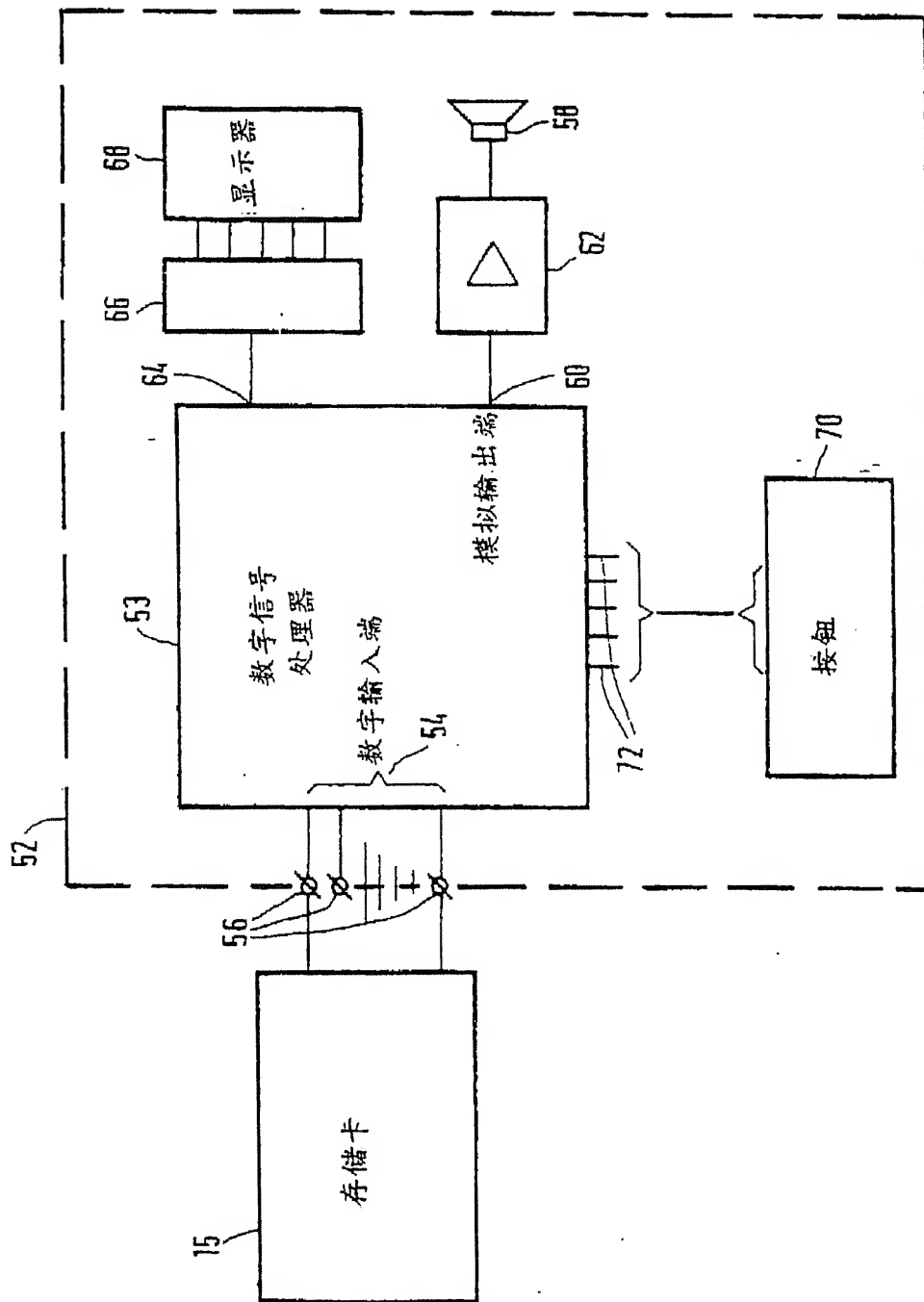


图 5